



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/532,533	03/22/2000	Michael James Taylor	1263.1240	7759

5514 7590 07/09/2004

FITZPATRICK CELLA HARPER & SCINTO
30 ROCKEFELLER PLAZA
NEW YORK, NY 10112

EXAMINER

ARMSTRONG, ANGELA A

ART UNIT PAPER NUMBER

2654

DATE MAILED: 07/09/2004

13

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary

Application No.

09/532,533

Applicant(s)

TAYLOR ET AL.

Examiner

Angela A. Armstrong

Art Unit

2654

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☒ Responsive to communication(s) filed on 16 March 2004.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 1-92 is/are pending in the application.
- 4a) Of the above claim(s) 43-86 is/are withdrawn from consideration.
- 5) ☐ Claim(s) _____ is/are allowed.
- 6) ☒ Claim(s) 1-42 and 87-92 is/are rejected.
- 7) ☐ Claim(s) _____ is/are objected to.
- 8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on _____ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☒ All b) ☐ Some * c) ☐ None of:
1. ☒ Certified copies of the priority documents have been received.
2. ☐ Certified copies of the priority documents have been received in Application No. _____.
3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- 1) ☒ Notice of References Cited (PTO-892)
- 2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
- 3) ☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)
Paper No(s)/Mail Date 3, 6, 7, 8, 10.
- 4) ☐ Interview Summary (PTO-413)
Paper No(s)/Mail Date. _____.
- 5) ☐ Notice of Informal Patent Application (PTO-152)
- 6) ☐ Other: _____.

DETAILED ACTION

Election/Restrictions

1. Applicant's election without traverse of embodiment I (claims 1-42) in Paper No. 11 is acknowledged.
2. Applicant is reminded that upon the cancellation of claims to a non-elected invention, the inventorship must be amended in compliance with 37 CFR 1.48(b) if one or more of the currently named inventors is no longer an inventor of at least one claim remaining in the application. Any amendment of inventorship must be accompanied by a request under 37 CFR 1.48(b) and by the fee required under 37 CFR 1.17(i).

Specification

3. The lengthy specification has not been checked to the extent necessary to determine the presence of all possible minor errors. Applicant's cooperation is requested in correcting any errors of which applicant may become aware in the specification.

Claim Objections

4. Claims 1, 13, and 16 are objected to because of the following informalities: The claim status of claims 1, 13, and 16, indicate the claims are previously presented, however the claims have single brackets used in indicating deleted matter. The Examiner assumes the bracket is a typographical error from the previous claim submission. Applicant is respectfully requested to delete the brackets from the body of claims 1, 13, and 16. Appropriate correction is required.

Claim Rejections - 35 USC § 102

The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(e) the invention was described in (1) an application for patent, published under section 122(b), by another filed in the United States before the invention by the applicant for patent or (2) a patent granted on an application for patent by another filed in the United States before the invention by the applicant for patent, except that an international application filed under the treaty defined in section 351(a) shall have the effects for purposes of this subsection of an application filed in the United States only if the international application designated the United States and was published under Article 21(2) of such treaty in the English language.

5. Claims 12-14, 16, 31-33, 35-40, 42, and 87-92 are rejected under 35 U.S.C. 102(e) as being anticipate by Potts et al (US Patent No. 6,593,956).

Potts discloses a system for locating an audio source in a video conferencing system.

6. Regarding claim 12, Potts discloses

an image processor for processing image data recorded by at least one camera showing the movements of a plurality of people to track each person in three dimensions (Figure 2, element 20, Figure 3, col. 6, lines 44-64; col. 8, line 34-col. 9, line 59);

a sound processor for processing sound data to determine the direction of arrival of the sound (Figure 2, element 20, Figure 3, col. 6, lines 44-64; col. 18, lines 19-23);

a speaker identifier for determining which of the people is speaking based on the result of the processing performed by said image processor and the result of the processing performed by said sound processor (col. 18, line 55 to col. 19, line 52).

7. Regarding claim 13, Potts discloses the image processor is arranged to track each person by processing the image data using camera calibration data defining the position and orientation of each camera from which image data is processed (col. 6, lines 35-42; col. 7, line 55-col. 8, line 7; col. 8, lines 50-59).

Art Unit: 2654

8. Regarding claim 14, Potts discloses the image processor is arranged to track each person by tracking each person's head (col. 7, line 55-col. 8, line 30).

9. Regarding claim 16, Potts discloses the speaker identifier is arranged to identify a person who is speaking in a given frame of the received image data using the results of the processing performed by said image processor and said sound processor for at least one other frame if the speaker cannot be identified using the results of the processing performed by said image processor and said sound processor for the given frame (col. 13, line 14-col. 15, line 65; col. 8, lines 60-66).

10. Regarding claim 37, as dependent upon claim 12, Potts discloses a storage device storing computer program instructions for programming a programmable processing apparatus to become configured as an apparatus at col. 6, line 65-col. 7, line 22.

11. Regarding claim 39, as dependent upon claim 12, Potts discloses a signal conveying computer program instructions for programming a programmable processing apparatus to become configured as an apparatus at col. 6, line 65-col. 7, line 22.

12. Regarding claims 42, 87-88, and 91-92, claims 42, 87-88, and 91-92 are apparatus claims similar in scope and content to apparatus claims 12-14, 16, 37, and 39, and are therefore rejected under similar rationale.

13. Regarding claims 31-33, 35-36, 38, 40, and 89-90, claims 31-33, 35, and 89-90 are method claims similar in scope and content to apparatus claims 12-14, 16, 37, and 39, and are therefore rejected under similar rationale.

Claim Rejections - 35 USC § 103

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

14. This application currently names joint inventors. In considering patentability of the claims under 35 U.S.C. 103(a), the examiner presumes that the subject matter of the various claims was commonly owned at the time any inventions covered therein were made absent any evidence to the contrary. Applicant is advised of the obligation under 37 CFR 1.56 to point out the inventor and invention dates of each claim that was not commonly owned at the time a later invention was made in order for the examiner to consider the applicability of 35 U.S.C. 103(c) and potential 35 U.S.C. 102(e), (f) or (g) prior art under 35 U.S.C. 103(a).

15. Claims 1-4, 6, 17-20, 22, 23, and 37-41 are rejected under 35 U.S.C. 103(a) as being unpatentable over Potts et al (US Patent No. 6,593,956) in view of Brais (US Patent No. 5,995,936).

16. Regarding claim 1, Potts discloses

an image processor for processing image data recorded by at least one camera showing the movements of a plurality of people to track each person in three dimensions (Figure 2, element 20, Figure 3, col. 6, lines 44-64; col. 8, line 34-col. 9, line 59);

a sound processor for processing sound data to determine the direction of arrival of the sound (Figure 2, element 20, Figure 3, col. 6, lines 44-64; col. 18, lines 19-23);

a speaker identifier for determining which of the people is speaking based on the result of the processing performed by said image processor and the result of the processing performed by said sound processor (col. 18, line 55 to col. 19, line 52).

Potts does not teach a voice recognition processor for processing the received sound data to generate text data therefrom in dependence upon the result of the processing performed by said speaker identifier.

Brais teaches a report generation system and method for capturing prose, audio and video by voice command and automatically linking sound and image to formatted text locations. Brais teaches the system can be used to generate reports in an environment of a plurality of users using a plurality of video cameras (col. 13, lines 2-7). At col. 9, lines 25-40, Brais teaches the system uses voice recognition or speech to text means to convert received signals into digital signals and interprets the digital signals as prose in the form of words, expressions, descriptions, or commands. Brais teaches the system allows for compilation of electronic reports with a single data-gathering step and a fully automated report compilation step (col. 4, lines 14-16).

Therefore, it would have been obvious to one of ordinary skill at the time of the invention to modify the system of Potts to implement a voice recognition processor, as suggested by Brais, for the purpose of providing an automatic electronic report from the image and audio data acquisition of a plurality of users and video cameras, as also suggested by Brais.

17. Regarding claim 2, Potts does not teach a voice recognition processor including a storage unit for storing voice recognition parameters for each of the plurality of people. Brais teaches the system can be used to generate reports in an environment of a plurality of users using a plurality of video cameras (col. 13, lines 2-7) and that the system supports speaker-dependent

Art Unit: 2654

speech recognition (col. 9, lines 9-12), which reads on “a voice recognition processor including a storage unit for storing voice recognition parameters for each of the plurality of people”, since a speaker-dependent system requires the speech characteristics of that particular individual.

It would have been obvious to one of ordinary skill at the time of the invention to modify the system of Potts to implement speaker-dependent speech recognition, as suggested by Brais, for the purpose of providing more accurate speech recognition of the audio data and thereby yielding the electronic report with minimal errors and corrections.

Potts and Brais do not teach a selection processor for selecting the voice recognition parameters. However, selecting a specific set of recognition parameters from a database of recognition parameters is well known in a multiple user speaker-dependent recognition system.

Therefore, it would have been obvious to one of ordinary skill at the time of the invention to modify the system of Potts to implement speaker-dependent speech recognition, as suggested by Brais, and to further implement a means of selecting the appropriate parameters for the specific individual producing the audio, as was well known in the art, for the purpose of increasing the accuracy and efficiency of the recognition processing.

18. Regarding claim 3, Potts discloses the image processor is arranged to track each person by processing the image data using camera calibration data defining the position and orientation of each camera from which image data is processed (col. 6, lines 35-42; col. 7, line 55-col. 8, line 7; col. 8, lines 50-59).

19. Regarding claim 4, Potts discloses the image processor is arranged to track each person by tracking each person's head (col. 7, line 55-col. 8, line 30).

Art Unit: 2654

20. Regarding claim 6, Potts discloses the speaker identifier is arranged to identify a person who is speaking in a given frame of the received image data using the results of the processing performed by said image processor and said sound processor for at least one other frame if the speaker cannot be identified using the results of the processing performed by said image processor and said sound processor for the given frame (col. 13, line 14-col. 15, line 65; col. 8, lines 60-66).

21. Regarding claim 37, as dependent upon claim 1, Potts discloses a storage device storing computer program instructions for programming a programmable processing apparatus to become configured as an apparatus at col. 6, line 65-col. 7, line 22.

22. Regarding claim 39, as dependent upon claim 1, Potts discloses a signal conveying computer program instructions for programming a programmable processing apparatus to become configured as an apparatus at col. 6, line 65-col. 7, line 22.

23. Regarding claim 41, claim 41 is an apparatus claims similar in scope and content to apparatus claims 1-4, 6, 37, and 39, and are therefore rejected under similar rationale.

24. Regarding claims 17-20, 22, 23, 38, and 40, claims 17-20, 22, 23, 38, and 40 are method claims similar in scope and content to apparatus claims 1-4, 6, 37, and 39 and are therefore rejected under similar rationale.

25. Claims 5, 7-11, 21, and 24-30 rejected under 35 U.S.C. 103(a) as being unpatentable over Potts in view of Brais as applied to claims 1 and 17 above, and further in view of Andersson et al (US Patent No. 5,500,671).

Art Unit: 2654

26. Regarding claim 5, Potts and Brais do not teach the image processor is arranged to process the image data to determine where at least each person who is speaking is looking.

Andersson teaches a video conference system which provides eye contact and a sense of presence to a plurality of conference participants. Andersson teaches eye gaze tracking processing for determining where a person is looking while they are speaking (col. 6, line 3 to col. 7, line 65) for the purpose of providing information on who is talking to whom during the video conference so as to provide the appearance of eye contact to preserve the look and feel of a live meeting (col. 7, lines 16-26).

Therefore, it would have been obvious to one of ordinary skill at the time of the invention to modify the system of Potts and Brais to provide eye gaze tracking and processing, as taught by Andersson, for the purpose of providing information on who is talking to whom during the video conference so as to provide the appearance of eye contact to preserve the look and feel of a live meeting when a plurality of conference participants are using the video conferencing system, as suggested by Andersson (col. 7, lines 16-26).

27. Regarding claim 7, Potts does not teach viewing data defining where at least each person who is speaking is looking.

Andersson teaches a video conference system which provides eye contact and a sense of presence to a plurality of conference participants. Andersson teaches eye gaze tracking processing for determining where a person is looking while they are speaking (col. 6, line 3 to col. 7, line 65) for the purpose of providing information on who is talking to whom during the video conference so as to provide the appearance of eye contact to preserve the look and feel of a live meeting (col. 7, lines 16-26).

Therefore, it would have been obvious to one of ordinary skill at the time of the invention to modify the system of Potts to provide eye gaze tracking and processing, as taught by Andersson, for the purpose of providing information on who is talking to whom during the video conference so as to provide the appearance of eye contact to preserve the look and feel of a live meeting when a plurality of conference participants are using the video conferencing system, as suggested by Andersson (col. 7, lines 16-26).

Potts does not teach a database for storing at least some of the received image data, the sound data, the text data produced by said voice recognition processor and viewing data defining where at least each person who is speaking is looking, said database being arranged to store the data such that corresponding text data and viewing data are associated with each other and with the corresponding image data and sound data.

Brais teaches associating the audio, image and text data and the storing of the associated data (col. 12, line 26 to col. 13, line 39; Figure 13), for later retrieval and/or during report generation.

It would have been obvious to one of ordinary skill at the time of the invention to modify the system of Potts to provide eye gaze tracking and processing, as taught by Andersson, for the purpose of providing information on who is talking to whom during the video conference so as to provide the appearance of eye contact to preserve the look and feel of a live meeting, as suggested by Andersson (col. 7, lines 16-26) and to further modify the system to provide for the association and storage of the audio, text, image, and eye gaze tracking data, for the purpose of having access to the data for later retrieval for generating reports of the conferencing activities, as taught by Brais.

Art Unit: 2654

28. Regarding claims 8 and 9, Potts and Brais do not teach a data compressor, which comprises a data encoder for encoding the image data and the sound data as MPEG data.

Andersson teaches encoding of the image data for bit rate reduction to facilitate transmission over a via a telecommunication network (col. 3, lines 56-67).

It would have been obvious to one of ordinary skill at the time of the invention to modify the system of Potts and Brais to implement data encoding of the image data, as taught by Andersson, for the purpose of achieving bit rate reduction of the data to facilitate transmission and/or storage of the video data.

Potts, Brais, and Andersson do not teach encoding the sound data as MPEG data. However, encoding sound data as MPEG data was well known in the art, and it would have been obvious to one of ordinary skill to modify the system of Potts, Brais and Andersson to provide for MPEG data encoding, as was well known in the art, for the purpose of achieving bit rate reduction of the data to facilitate transmission and/or storage of the audio data.

29. Regarding claims 10 and 11, Potts does not teach a gaze data generator for generating data defining, for a predetermined period, the proportion of time spent by a given person looking at each of the other people during the predetermined period and wherein the predetermined period comprises a period during which the given person was talking.

Andersson teaches a video conference system which provides eye contact and a sense of presence to a plurality of conference participants. Andersson teaches eye gaze tracking processing for determining where a person is looking while they are speaking, which reads on "the predetermined time period comprises a period during which the given person was talking", (col. 6, line 3 to col. 7, line 65) for the purpose of providing information on who is talking to

Art Unit: 2654

whom during the video conference so as to provide the appearance of eye contact to preserve the look and feel of a live meeting (col. 7, lines 16-26).

Therefore, it would have been obvious to one of ordinary skill at the time of the invention to modify the system of Potts to provide eye gaze tracking and processing, as taught by Andersson, for the purpose of providing information on who is talking to whom during the video conference so as to provide the appearance of eye contact to preserve the look and feel of a live meeting, as suggested by Andersson (col. 7, lines 16-26).

Potts does not teach the database is arranged to store the data so that it is associated with the corresponding image data, sound data, text data and viewing data.

Brais teaches associating the audio, image and text data and the storing of the associated data (col. 12, line 26 to col. 13, line 39; Figure 13), for later retrieval and/or during report generation.

It would have been obvious to one of ordinary skill at the time of the invention to modify the system of Potts to provide eye gaze tracking and processing, as taught by Andersson, for the purpose of providing information on who is talking to whom during the video conference so as to provide the appearance of eye contact to preserve the look and feel of a live meeting, as suggested by Andersson (col. 7, lines 16-26) and to further modify the system to provide for the association and storage of the audio, text, image, and eye gaze tracking data, for the purpose of having access to the data for later retrieval for generating reports of the conferencing activities, as taught by Brais.

30. Regarding claims 21 and 24-30, claims 21 and 24-30 are method claims similar in scope and content to apparatus claims 5, and 7-11 and are therefore rejected under similar rationale.

31. Claims 15 and 34 are rejected under 35 U.S.C. 103(a) as being unpatentable over Potts in view of Andersson et al (US Patent No. 5,500,671).

32. Regarding claims 15 and 34, Potts does not teach the image processor is arranged to process the image data to determine where at least each person who is speaking is looking.

Andersson teaches a video conference system which provides eye contact and a sense of presence to a plurality of conference participants. Andersson teaches eye gaze tracking processing for determining where a person is looking while they are speaking (col. 6, line 3 to col. 7, line 65) for the purpose of providing information on who is talking to whom during the video conference so as to provide the appearance of eye contact to preserve the look and feel of a live meeting (col. 7, lines 16-26).

Therefore, it would have been obvious to one of ordinary skill at the time of the invention to modify the system of Potts to provide eye gaze tracking and processing, as taught by Andersson, for the purpose of providing information on who is talking to whom during the video conference so as to provide the appearance of eye contact to preserve the look and feel of a live meeting when a plurality of conference participants are using the video conferencing system, as suggested by Andersson (col. 7, lines 16-26).

Art Unit: 2654

Conclusion

33. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Angela A. Armstrong whose telephone number is 703-308-6258. The examiner can normally be reached on Monday-Thursday 7:30-5:00 PM.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Richemond Dorvil can be reached on (703) 305-9645. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

Angela A. Armstrong
Examiner
Art Unit 2654

AAA
June 24, 2004

Angela Armstrong